An overlapping site model for the lipid annulae of membrane proteins

A.G. Lee

Department of Biochemistry, University of Southampton, Southampton, SO9 3TU, England

Received 1 November 1982

A model is presented for the binding of phospholipids and related molecules to membrane proteins, based on a fixed number of binding sites for fatty acyl chains. It is shown that in such a model, statistical factors favour binding of single chain molecules such as fatty acids when present at low concentrations in mixtures with phospholipids

Membrane proteins *Lipid-protein interactions* *Fatty acids* *Cardiolipin*

1. INTRODUCTION

Both electron spin resonance (ESR) and fluorescence techniques have been used to study the interactions of phospholipids and other hydrophobic molecules with membrane proteins and to obtain relative binding constants at the lipid–protein interface [1–8]. But it is unclear how the binding of phospholipids containing two fatty acyl chains (such as phosphatidylcholines) should be compared to the binding of a single chain molecule such as a fatty acid or to a lipid such as cardiolipin containing 4 fatty acyl chains. Previously it has been assumed that all these molecules occupy a single site on the protein surface, but if the area of the lipid–protein interface is to be conserved, this seems intrinsically unlikely. A more reasonable model would be one in which a fixed number of fatty acyl chains are necessary to cover the hydrophobic protein surface so that fatty acids, phosphatidylcholines and cardiolipins will occupy, respectively, 1, 2 and 4 sites around the protein (fig.1). For such a model, binding can no longer be described by simple Scatchard analysis, for the following reasons. Consider an initial state in which all sites are occupied by fatty acids. If now a phospholipid is added that occupies two sites on the protein surface then the situation illustrated in fig.1 could develop. Random binding of the phospholipid could result in a single site being left unoccupied between two bound phospholipid molecules (as at A). Again, if binding of phospholipid molecules creates 3 consecutive sites (as at B), then it is only possible for one phospholipid molecule to bind at these 3 sites. For statistical reasons of this kind, it becomes difficult to completely saturate the sites with a multivalent ligand such as a phospholipid.

An exactly analogous problem arises in the study of ligand binding to pairs of sites on DNA, and has been analysed in [9]. We show here that the method of McGhee and von Hippel can be applied to our problem.

2. THE MODEL

We make a number of simplifying assumptions:

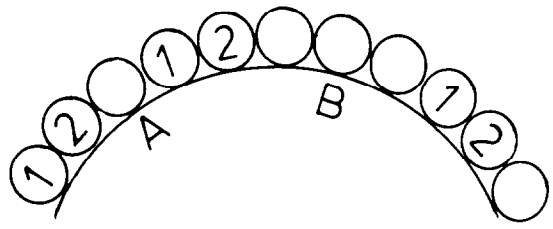

Fig.1. Binding sites on the surface of a membrane protein, showing 3 phospholipids bound at random, indicating the sites occupied by the 1- and 2- fatty acyl chains.

Published by Elsevier Biomedical Press
00145793/83/0000–0000/$3.00

(i) The lipid–protein interface consists of a fixed number N of binding sites for fatty acyl chains, of equal affinity.
(ii) The number of binding sites (N) is considerably greater than the number of adjacent sites (n) occupied by a particular ligand molecule. The system thus reduces to a linear lattice of N identical repeating binding sites in which end-effects can be ignored.
(iii) Partial binding of a ligand is not allowed. Thus, for example, a phosphatidylcholine ($n = 2$) must occupy two fatty acyl sites. This assumption is justified below, but can be relaxed.
(iv) Binding is non-cooperative. Cooperativity can be readily introduced [9], but there is insufficient experimental data to make it presently worthwhile.

We consider a protein in a membrane composed of a mixture of monovalent ligand L (e.g., fatty acid) and multivalent ligand L^* (e.g., phosphatidylcholine). Binding to the surface of the protein is described by a series of displacement reactions of the type:

$$PL_n + L^* \longrightarrow PL^* + nL$$

for which we can write an equilibrium constant K,

$$K = [PL^*][L_f]^n/[PL_n][L_f^*] \quad (1)$$

where:
square brackets denote concentrations;
$[L_f]$ and $[L_f^*]$ are the concentrations of unbound L and L^* respectively. For convenience, we have used a nomenclature similar to that in [4] where it is shown that, when all L and L^* are in the membrane (which will be true for very hydrophobic molecules), the natural unit for concentration and standard state is the mole fraction.

The equilibrium constant can be rewritten in terms of the more usual concentration units of mol of ligand/l of medium:

$$K = nL_b^*L_f^n/L_bL_f^*(L_f + L_f^*)^{n-1} \quad (2)$$

where:
L_b and L_b^* are the concentrations (mol/l) of L and L^*, respectively bound to the protein;
L_f and L_f^* are the corresponding unbound concentrations.

From the various conservation relationships we have:
binding sites:

$$L_b + nL_b^* = NP \quad (3)$$

monovalent ligand:

$$L_f + L_b = L_t \quad (4)$$

multivalent ligand:

$$L_f^* + L_b^* = L_t^* \quad (5)$$

where P is the protein concentration.

It has been shown [9] that in the description of the binding of a multivalent ligand, the free site concentration has to be multiplied by a factor P_n, giving the probability that a free site will be followed by at least $n - 1$ other free sites, thus allowing binding of a ligand of valence n. In our case, equations 2–5 can be readily combined to give an equation analogous to the Scatchard equation:

$$L_b^* = K\left\{\frac{L_f^*(L_f + L_f^*)^{n-1}}{nL_f^n}\right\}(NP - nL_b^*)P_n \quad (6)$$

Simple statistical arguments [9] give the probability factor P_n as

$$P_n = \{(NP - nL_b^*)/(NP - n - 1)L_b^*)\}^{n-1} \quad (7)$$

Equation (6) assumes 'stereochemical' binding of the multivalent ligand: that is, the ligand can bind only one way round to the site. If this requirement is relaxed, then it can be shown that P_n is unaltered, but that K in equation (6) should be replaced by $2K$ to allow for the two statistical ways of binding. For monovalent ligands, P_n reduces to one, and equation (6) has the form of the simple Scatchard equation as derived in [4]. For multivalent ligands, P_n is always less than unity, and as it is a function of L_b^* it will give curved binding plots. The effect of the P_n term becomes large when L_b^*/N is large.

Unfortunately equation (6) cannot be written in a form which allows direct calculation of L_b^*. However, the equation can be readily solved for L_b^* using numerical methods, such as the method of bisection [10]. Again following [4] we define:

$$x = L_t/P,\ x^* = L_t^*/P;$$

$$y = L_f/(L_t - L_f);$$

$$y^* = L_f^*/(L_t^* - L_f^*).$$

Fig. 2 and 3 compare calculated binding plots for mixtures of ligand of the same valence with those for mixtures of monovalent and divalent ligand. Calculated ratios of bilayer/bound divalent ligand in mixtures with a monovalent ligand are similar to calculated ratios of bilayer/bound monovalent ligand in mixtures of two monovalent ligand, when the relative binding constant K for the former is twice that of the latter (fig.2). Differences between the two sets of binding curves are greatest at high mol fractions of the divalent ligand, as expected. The two sets of curves become identical over intermediate concentration ranges if the relative concentrations of ligands are expressed on the basis of fatty acyl chain concentration (i.e., 1 phospholipid = 2 fatty acyl chains), and the factor of

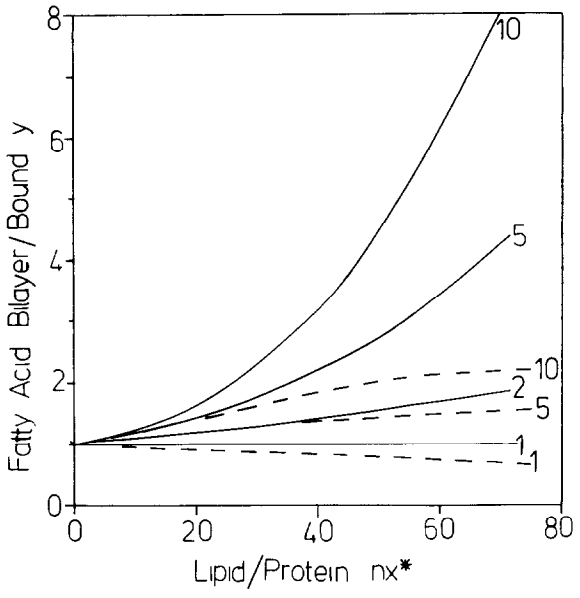

Fig.2. Calculated ratios of bilayer/bound concentrations for a ligand L^* in mixtures with a second ligand L, as a function of the ratio of the ligand L^* to protein concentration, expressed as nx^* where n is the number of sites occupied by the ligand L^*. The number of binding/protein, $N = 40$, and the ratio of the total concentrations of L^* and L to protein concentration is maintained at $x + nx^* = 80$. Each curve corresponds to the marked value of the relative binding constant K. Solid lines correspond to the case where both L^* and L are monovalent ligands, so that $n = 1$. Broken lines correspond to the case where L^* is a divalent ligand and L is a monovalent ligand, so that $n = 2$. The calculated curve for $K = 2$ and $n = 2$ is identical to that shown with $K = 1$ and $n = 1$.

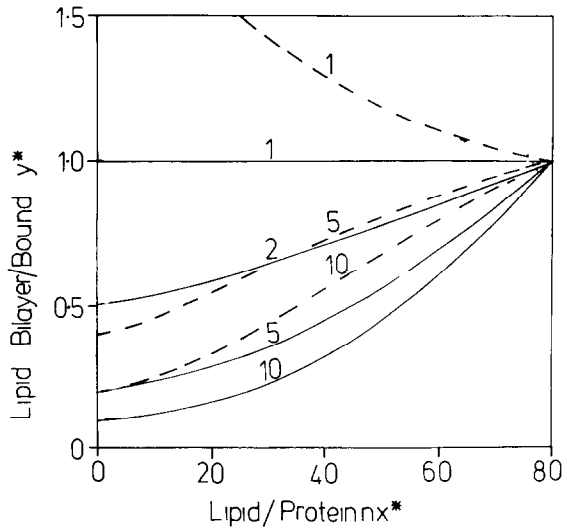

Fig.3. Calculated ratios of bilayer/bound concentrations for a ligand L in mixtures with a second ligand L^*, under the conditions given in the legend to fig.2.

2 is introduced into the relative binding constant. Binding curves for monovalent ligand in mixtures with another monovalent ligand are very different from those for mixtures with a divalent ligand when the mixtures contain a high mol fraction of divalent ligand (fig.3). The curves reflect the difficulty in saturating the sites around the protein with divalent ligand, because of the statistical factor (equation (7)). This point is also clearly made in fig.4, which represents an experiment in which the concentration of monovalent ligand is kept constant at a low value, and the concentration of divalent ligand is increased.

The overlapping site model presented here gives rise to a rather different interpretation of ESR data on the binding of spin-labelled compounds to membrane proteins. Table 1 illustrates this with some simulated data based on experiments with rhodospin [2]. Typically in ESR experiments, the spin-labelled derivative is present as a small fraction (1%) of the total lipid, to prevent line-broadening effects. Experiment A of table 1 represents a measurement of the bilayer/bound ratio of a spin-labelled fatty acid in the presence of a two-chain phospholipid. If it is assumed that the fatty acid and phospholipid compete in a one-to-one fashion for sites on the protein (so that for both ligands $n = 1$) then, as expected, the bilayer/bound

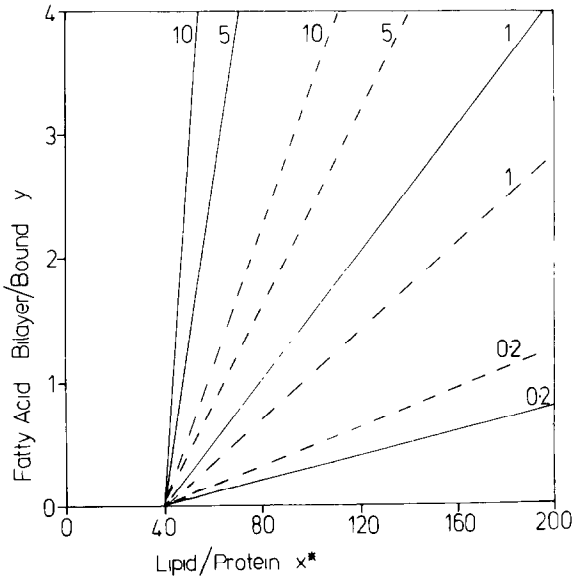

Fig.4. Calculated ratio of bilayer/bound concentrations for a fatty acid in mixtures with a phospholipid ($n = 2$), as a function of the phospholipid/protein ratio x^*, for the marked values of the relative binding constants. The fatty acid/protein ratio was fixed at 1.0. Solid lines represent calculations based on a one-to-one displacement model, with the number of binding sites/protein $N = 40$. Broken lines represent calculations based on the overlapping site model, with the number of fatty acyl chain sites/protein $N = 80$.

ratio for the fatty acid is linearly related to the binding constant for the phospholipid relative to that for the fatty acid (K). This simple relationship no longer holds if the phospholipid is assumed to occupy two sites on the protein, because statistical factors increase the probability that at least a few sites on the protein will be occupied by fatty acid. The problem no longer arises if the experiment is done the other way round, with the spin-labelled derivative being the multivalent ligand (experiment B, table 1). Binding constants derived from the two models simply differ by a factor of close to two, reflecting the requirement for two monovalent ligands to displace one divalent ligand in the overlapping site model.

The model presented here can be used to reinterpret some published ESR experiments, as shown in table 2. For the experiments with cardiolipin spin labels, the binding constants in table 2 are about twice those calculated on the basis of one-to-one

Table 1

Calculated bilayer/bound molar ratios for a protein with $N = 24$, in which one ligand (ligand 1) is monovalent and the other is either monovalent or divalent

	Molar ratio Ligand 1 protein	Molar ratio Ligand 2 protein	Ligand 2 K[a]		Ligand 1 Bilayer bound
			$n = 1$	$n = 2$	
A	0.65	65.0	0.6	0.7	1.0
	0.65	65.0	1.15	2.8	2.0
	0.65	65.0	1.75	6.0	3.0
					Ligand 2 Bilayer bound
B	65.0	0.65	1.67	0.70	1.0
	65.0	0.65	0.87	0.35	2.0
	65.0	0.65	0.57	0.23	3.0

[a] Binding constant for ligand 2 relative to ligand 1

displacement. However, there is no simple relationship between the binding constants for fatty acid spin labels calculated on the basis of the two models. The calculated binding constants are insensitive to the assumed spin label concentration in this region.

3. DISCUSSION

Binding of phospholipids and fatty acyl derivatives to membrane proteins has been discussed in terms of a fixed number of fatty acyl chain binding sites on the hydrophobic surfaces of the protein. A statistical approach allows the calculation of the composition of the annulae around membrane proteins in the presence of mixtures of hydrophobic derivatives of different valence, such as fatty acids ($n = 1$) phosphatidylcholines ($n = 2$) and cardiolipin ($n = 4$). The model predicts that in mixtures of phospholipid containing a low concentration of fatty acid, a relatively large fraction of the fatty acid will be bound to the protein, for statistical reasons. Absolute values of protein affinities for spin-labelled fatty acids made using ESR methods, will therefore tend to be overestimates. It is not possible to use high concentrations of spin label where statistical effects are less important because of line broadening effects. Measurements of relative binding constants can, however, be made over

Table 2

Binding parameters estimated from spin-label experiments

Protein	Ligand[a]	N^b	x^c	x^{*c}	$Y_{expt}{}^b$	$Y^*_{expt}{}^b$	$K_{rel}{}^d$ divalent/ monovalent
Rhodopsin	14-SASL	48	0.65	65.0	1.94	–	2.5
	14-CLSL	24	65.0	0.65	–	2.57	1.4
Na/K-ATPase	14-SASL	116	3.0	305.0	2.57	–	0.6
	14-CLSL	58	305.0	3.0	–	1.0	9.0

[a] SASL, stearic acid spin label; CLSL, cardiolipin spin label
[b] Taken from [2,3]
[c] Calculated assuming a spin label concentration of 1% of the total lipid concentration
[d] Calculated assuming stereospecific binding

a wider concentration range using fluorescence quenching techniques, and elsewhere we have shown that binding constants for fatty acids to the $(Ca^{2+}-Mg^{2+})$-ATPase can be obtained in this way [7].

In the above, analysis, the possibility of partial binding of the ligand to the protein surface has been discounted. Thus it has been assumed that when a phosphatidylcholine binds to the protein, both fatty acyl chains must occupy sites on the protein, and binding of only one chain is now allowed. There are two reasons for thinking the assumption sound. Firstly, on binding to the protein surface, a molecule will lose much entropy (translational, rotational and conformational), and most of this loss will occur after the initial binding event. Binding of both fatty acyl chains of a phospholipid will therefore be favoured over binding of a single chain. Secondly, the unbound fatty acyl chain of a phospholipid attached to a protein by a single chain would be expected to have a mobility intermediate between that for a phospholipid fully bound to a protein and a phospholipid in a bilayer: no evidence for such an intermediate state has been reported [1–4].

Finally, our analysis has assumed a constant number of annular sites. That is, we have discounted the possibility of random protein–protein contacts which would result in a number of annular sites that would vary with protein concentration in the membrane. This might seem to be at variance with the generally low selectivity of the annular sites for phospholipids [2–6], since low selectivity suggests weak binding. However, the most important factor in determining the distribution of molecules in the membrane is the relative energies of protein–phospholipid and protein–protein interactions in the membrane [11]. If random protein–protein interactions are very unfavourable, then proteins will maintain a constant phospholipid annulus even if the protein–phospholipid interaction is weak. Specific protein–protein interactions may, of course, be strong, leading to specific oligomerisation, with a phospholipid annulus then surrounding the oligomer.

REFERENCES

[1] Griffith, O.H. and Jost, P.C. (1979) in: Cytochrome Oxidase (King, T.E., ed) Elsevier Biomedical, Amsterdam, New York.
[2] Watts, A., Volotovski, I.D. and Marsh, D. (1979) Biochemistry 18, 5006–5013.
[3] Marsh, D., Watts, A., Pates, R.D., Uhl, R., Knowles, P.F. and Esmann, M. (1982) Biophys. J. 37, 265–271.
[4] Brotherus, J.R., Griffith, O.H., Brotherus, M.O., Jost, P.C., Silvius, J.R. and Hokin, L.E. (1981) Biochemistry 20, 5261–5267.
[5] London, E. and Feigenson, G.W. (1981) Biochemistry 20, 1939–1948.
[6] East, J.M. and Lee, A.G. (1981) Biochemistry 21, 4144–4151.
[7] Simmonds, A.C., East, J.M., Jones, O.T., Rooney, E.K., McWhirter, J. and Lee, A.G. (1983) Biochim. Biophys. Acta, in press.
[8] Moules, I.K., Rooney, E.K. and Lee, A.G. (1982) FEBS Lett. 138, 95–100.

[9] McGhee, J.D. and von Hippel, P.H. (1974) J. Mol. Biol. 86, 469–489.
[10] McCormick, J.M. and Salvadori, M.G. (1964) 'Numerical Methods in Fortran', Prentice-Hall Inc., New Jersey.
[11] Lee, A.G. (1978) Biochim. Biophys. Acta 507, 433–444.